

DOCUMENT RESUME

ED 399 590

CS 509 334

AUTHOR Geske, Joel
TITLE Legibility of Sans Serif Type for Use as Body Copy in Computer Mediated Communication.
PUB DATE Aug 96
NOTE 24p.; Paper presented at the Annual Meeting of the Association for Education in Journalism and Mass Communication (79th, Anaheim, CA, August 9-13, 1996).
PUB TYPE Speeches/Conference Papers (150) -- Reports -- Research/Technical (143)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Students; *Computer Mediated Communication; Higher Education; *Readability
IDENTIFIERS Helvetica Type; *Legibility; *Typeface; Visual Communication

ABSTRACT

A study examined type legibility on computer screens to determine type styles and type sizes that are most legible. Subjects, 107 college students, tested legibility of Helvetica type in 3 different sizes and 3 different faces. Results indicated no significant differences for legibility of Helvetica 12, 10, and 9 point type in the normal typeface at the resolution of the computer screens. Helvetica Bold significantly increases the legibility in most cases and at least does not decrease legibility. Helvetica Italics should be used with extreme caution or avoided. (Contains 9 references and 15 tables of data. An appendix presents each type face reproduced at screen resolution and enlarged 200%.) (Author/RS)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

Legibility of Sans Serif Type for Use as Body Copy in Computer Mediated Communication

By Professor Joel Geske
126 Hamilton Hall
Iowa State University
Ames IA 50011
geske@iastate.edu (515) 294-0477

Presented August 1996
AEJMC Visual Communication Division

Abstract

This experimental design used 107 subjects to test legibility of Helvetica type in three different sizes and 3 different faces.

The study found there are no significant differences for legibility of Helvetica 12, 10 and 9 point type in the normal face at screen resolutions. Helvetica Bold significantly increases the legibility in most cases and at the least does not decrease the legibility. Helvetica Italics should be used with extreme caution or avoided.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official DERI position or policy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY

J. Geske

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

Legibility of Sans Serif Type for Use as Body Copy in Computer Mediated Communication

Numerous studies have been done determining both legibility and readability of type in printed materials in mass communication settings (Tinker, 1963, McVey & Weigeshaus, 1973 and Felici) . However, a great many of the classic studies in readability and legibility are from the 1920-1940 era. Starting in the 1980's with the advent of the personal computer and in particular in the mid-1990's with the increased usage of the Internet and other computer mediated sources for transferring large amounts of information it is necessary to see if many of the basic rules of legibility and readability of type hold true for computer monitors.

Review of Literature

The terms legibility and readability have quite distinct meanings although they are inter-related and sometimes used interchangeably. Prior to 1940 the term "legibility" was used to discuss factors of ease and speed of reading. After 1940 some researchers began using the term "readability" as a broader and perhaps more meaningful term with legibility referring to the recognition of letter forms. However, with the advent of readability formulas for measuring the difficulty of reading material there has been some confusion. For the purposes of this study, the term legibility will be used to measure the human eye's ability to discern characters and words--to actually be able to correctly recognize the letter and word forms. Readability will be reserved for measures testing the ease of reading and understanding material (Tinker, 1963).

Basic notions about text and readability have developed over many centuries. One of the first recorded legibility tests was done in France in the 1790's. However, two major discoveries occurred around 1900 are among the most important concerning legibility and readability:

- 1) Experienced (not beginning) readers read in whole words, not a character at a time and,

- 2) Readers use saccadic leaps or jumps along a line of text and pause to read at regular intervals. Readers tend to read several words or a phrase at one time before the eye moves to the next grouping of words.

In a very fundamental way, these discoveries helped shift research from legibility of single characters to readability of words and phrases.

Some legibility problems are inherent in the design of the alphabet itself. In fact, some of the most used letters in the English alphabet are easy to confuse, including C - G; H - N; E - F in the upper case and c - e (along with a and o in some faces); b - d; and p - q in the lower case. On computer screens i - l - I - 1 (small i, small l, capital i and numeral one) can also cause confusion.

Many type and design books suggest rules for legibility. Serif type faces are generally considered more legible than sans-serif type faces due to the added information to the eye that the serif provides as well as the "line" the serifs provide to guide the eye. However, studies cited by Tinker in 1963 on studies conducted by Tinker and Paterson showed no statistically significant differences in readability of ten serif type faces. Two other studies in the 1960's tested sans vs. serif faces and showed no significant reading problems for the sans serif types (Tinker, 1996). At best, the research suggests that serif type faces may be more legible, but other studies indicate that properly typeset and formatted, sans serif faces do not pose serious legibility problems.

The computer screen monitor, however, brings additional problems to the issue of legibility of type. When graphics based fonts first started appearing in graphical interfaces such as Macintosh and later Windows applications, they were almost always stored and displayed in bitmap format. A typical computer monitor has 72 pixel resolution (or 72 pixels by 72 pixels in each square inch of screen) for Macintosh and 96 pixel for most PC's. The minimum stroke width is one pixel wide or high. The next possible stroke width is double that or two pixels wide. Especially in smaller size fonts (such as those used for body copy) this allows for very little leeway in displaying fonts. Bitmapped fonts can be enlarged, but only by multiplying the existing pixels which generally results in the "jaggies" or jagged edges on the enlarged

type. Likewise, type can be reduced, but only with the problem of open areas in some letters filling in or some important pixels (such as the cross arm of an A) being deleted. Therefore, most bitmap fonts are designed for a particular size even though this limits type size choices.

To compound the problem, when text is highlighted and the bold function is used to create bold face type, the computer generally doubles the vertical stroke width of the letter but leaves the horizontal stroke of the letter at the original height. This helps solve the problem of letters filling by not adding additional pixels on the horizontal stroke, but alters the character and design of the type on a fundamental level. Doubling the vertical pixels also has the effect of extending the serifs and causing them to run together. Even sans-serif types can have a problem with letters crowding or running together in a bold face.

"Italic" type also has inherent problems on screen. In fact, rather than true italic, most monitors actually display an oblique or slanted form of the standard face. Again, especially in smaller font sizes that are using letter strokes only one pixel wide, this slanting letter can take on a very jagged and difficult to read appearance.

More recently, outline font technology has come to the personal computer. Here the letter is outlined using a mathematical formula to create straight lines and curves. Once the outline is complete it is filled in solid. One problem, however, is that this method is much slower to produce a page full of text than bitmapped fonts. To solve this problem, once the font has been formed, it is converted to a bitmap form and saved or "cached" in the computer memory. This conversion is known as rasterization. Conversion takes some time, but once completed the display fonts can be as fast as the display of bitmapped fonts. Unfortunately, in the rasterization process some rounding off of numbers occurs. Some letters, such as an H, which should have two vertical strokes of about the same width may end up with one a pixel larger. On large display type or when using a 300 dot per inch printer it may be hardly noticeable. On screen, however, with only a 72 or 96 dot resolution and when using small type faces the problem can be dramatic and severely affect both legibility and readability (Petzgold).

In addition, while outline and advanced technologies like Adobe Type Manager (ATM) TrueType or Speedo use outlines and gray scales to smooth out letters they also cause additional problems. These technologies work well with larger type, but readers spend relatively little time on headlines and much more on body copy. For body copy, the difficulty in reading comes from fuzzy edges and making out characters. Gray scale and anti-aliased type increase the "fuzziness" and the eye has to battle to draw a sharp focus (Felici, 1996).

Sans serif type faces, with less difference in stroke widths may have less problems. One type designer (Chuck Bigelow, designer of the Lucinda Family or type) recommends sans serif because "when printed, the serifs on typefaces are only a tiny percentage of the typeface design. But on-screen, in order to display the serifs using the limited number of available pixels, they take up a much bigger proportion of the information than they do in print. Serifs should be small things--but on screen they become big...noise or distracting chunks of interference (Will-Harris, 1996).

It can be concluded that computer screen monitors present unique challenges to the typographer and designer. Many of the assumptions made for print legibility and readability may not hold true for on-screen type. Even ignoring the fact of a computer screen being luminous while traditional type studies were done on reflective print media, the very nature of type construction and display on the screen bring into question legibility factors. Simple legibility factors must be considered before serious research can be conducted on readability.

The Research Question

As more traditional media look to computer mediated delivery systems, the question of how to provide large amounts of information in the most legible, readable form becomes an important factor. Today, with the increasing popularity of the Internet, we find books, magazines, newspapers, advertising and even videos on the computer. The fundamental question is to look at type legibility on screen to determine type styles and type sizes that are most legible.

The review of literature and the exploration of problems encountered with screen resolutions does not allow one to draw conclusions based on previous research concerning serif vs. sans serif faces. While previous research in print suggests serif faces are easier to read, problems with screen resolution, especially when making faces bold, would suggest that sans-serif faces may be the better choice for screen presentation. Serif type faces tend to use more thick/thin differentiation in strokes that can lead to rasterization problems and increased problems with bowls (spaces in letters such as "e") filling in. Finally, serifs can run together, especially on bold type. Because of these potential problems with serif type, a sans serif type face will be explored initially for a study on the effects of the variables of size, boldness and italics on legibility of a type face.

This study looks at three basic research questions:

- 1) What is the most legible type size for body copy or long segments of text?
- 2) Does bold facing enhance or reduce legibility at different type sizes?
- 3) Do italic faces affect legibility?

Methodology

Tinker provides a comprehensive discussion of methodologies for testing legibility. This study will use the Speed of Perception method for testing character legibility. By using a very short exposure technique, the quickness and accuracy in perception of words, symbols and phrases can be measured. This method has been used successfully in previous research to study factors such as use of serifs, boldness of letters and optimal width stroke and optimal height/width ratios of letters (Tinker, 1963).

In this method, the human subject observes a pre-exposure field, is briefly exposed to the message and ends with a post-exposure field. The exposure to the message is very short, generally in the range of 1/10 of one second to 1/100 of a second.

For this experiment, the subject is seated in front of a computer screen. The screen has a black background with a white rectangular field in the upper third of the screen at approximately eye level. Below the rectangular field is a button for the subject to press when (s)he is ready for the exposure. The computer is programmed to provide 1/60 of one second exposure before moving to the post exposure field to block the words. This time of exposure was pre-tested to find a time that provided differentiation between type styles and sizes. This exposure time yields a single act of vision since the timing prevents a second fixation or saccadic leap. Older studies used a tachistoscope to view type printed on cards. This experiment uses the actual computer screen for the subjects to view.

For this experiment the researchers used black type on a white background. This is considered one of the most legible combinations in print studies. While research suggests black on white may not be the most readable for long text on computer screens, this experiment uses short phrases and short exposures and each type face and style is being viewed with the same color combination.

Helvetica was selected as a sans serif type face as it is commonly available Macintosh computers (Ariel is the equivalent on many PC versions).

Three body type sizes, 9 point, 10 point and 12 point, were selected. Sizes below 9 point were considered, however, when type set in sizes smaller than 9 point are magnified, one can clearly see that many letters lose all definition and simply become square blocks of pixels. These sizes were chosen because as more documents go electronic and are read and formatted by such programs such as Acrobat or Common Ground there will likely be more and more pressure to design in computer screen proportions (or what would be considered half-pages in print). This design puts pressure on designers to keep sizes relatively small in the 10-12 point range (Felici, 1996) Sizes over 12 point were not considered efficient for body copy.

Each type size was tested in normal, bold and italic.

This arrangement yielded 9 cells in the trial. For each exposure a phrase was inserted in the exposure field. Each phrase consisted of three common words that form a simple sentence such as "Trees have leaves" or "Dogs chase cats". Phrases were chosen as most adults take in several words in a saccadic leap or fixation. Each phrase had 15, 16 or 17 (16 plus or minus one) characters and spaces.

Phrases were randomly assigned to type faces and styles and each type style was randomly selected for order of testing. To help control for difficulty of reading a particular phrase, two forms were tested with a second random assignment of phrases to type face. No phrase and type face combinations were repeated from Form A to Form B.

Since speed of character recognition is essential, subjects were screened and eliminated for uncorrected vision impairments, learning disabilities that might impair reading or if their primary language was not English.

Subjects were given several examples to adjust to the procedure and methodology before the test began. Subjects controlled the rate of testing by using the mouse to click on a button directly below the exposure area. Subjects recorded each word in the phrase that they could perceive on a form that could be checked for accuracy. Each word had to be exact to be counted.

Subjects chosen were students at a major midwestern university in the United States. College students were chosen for several reasons. First, one of the largest and fastest growing areas of computer mediated communication is the Internet. One of the most comprehensive and current surveys on Internet use still shows students as making up over half of the total users (even with the increased number of commercial users) (Matrix News, 1995). College students in general are literate, familiar with computers and are at an age where there are minimal uncorrected vision problems. This should be viewed as an optimal audience, however, with differences expected with less literate subjects or at different age groups.

A sample of 122 students were used in the experiment. Students were volunteers from a Principles of Advertising class with over 10 majors

represented in areas of business, communication, design, human performance and consumer sciences. Results from 15 respondents were not used because they had learning disabilities that interfered with reading or English was not their primary language. This left a sample of 107.

Mean age was 21.3 with: 3 subjects age 18; 12 subjects age 19; 36 subjects age 20; 35 subjects age 21; 8 subjects age 22; 5 subjects age 23; 2 subjects age 24; 4 subjects age 25; 1 subject age 26 and 1 subject age 32.

Gender for the experiment was 56% female and 44% male or just about flipped from the educational users of the Internet which are about 59% male and 41% female.

Findings for Helvetica

The legibility tests for Helvetica were run with two test groups. One concern was that legibility could be affected by the phrase chosen rather than type characteristics. Therefore, each type size and style was replicated using a different randomly selected phrase. A significance level of $p=0.01$ was chosen as a conservative measure to make sure that groups not statistically different due to phrase chosen. Of the nine cells, seven showed no significant difference at the 0.01 level, as shown in Table 1.

Table 1	n=54	n=53		
Type Style	Sample A Mean	Sample B Mean	T-test p value	Significant at 0.01 level
Helvetica 12	1.72	1.83	$p=0.570$	no
Helvetica 10	1.70	1.91	$p=0.380$	no
Helvetica 9	1.74	2.08	$p=0.092$	no
Helvetica 12 bold	2.13	2.21	$p=0.65$	no
Helvetica 10 bold	2.11	1.83	$p=0.11$	no
Helvetica 9 bold	1.79	1.81	$p=0.92$	no
Helvetica 12 italic	0.93	1.83	$p=0.000$	yes
Helvetica 10 italic	0.61	1.62	$p=0.000$	yes
Helvetica 9 italic	0.46	0.87	$p=0.015$	no

Overall, there appears to be no significant differences in Sample A and Sample B for normal and bold type faces. However, it does appear that the

words or phrases may make a difference in the more difficult to read italic faces. The italic faces as a group were less legible and the researcher will address this issue in the discussion section.

Having two samples adds strength to the results as it provides an initial study and a replication to confirm reliability of the measurement and validity for finding of differences between type sizes and styles. However, caution should be used when data is pooled and compared for the italic type faces.

Given that the two sample groups are roughly equal and that Sample A and Sample B are not statistically different in most cases, the question of differences in type style and size can be addressed. When comparing type the researcher was looking for broader trends and differences between sizes and styles. Here a significance level of $p=0.05$ was used.

Comparison of Helvetica Type by Size

Null Hypotheses (NH)

- NH1: There is no significant difference between Helvetica 12 point and Helvetica 10 point type viewed on computer monitors.
- NH2: There is no significant difference between Helvetica 10 point and Helvetica 9 point type viewed on computer monitors.
- NH3: There is no significant difference between Helvetica 12 point and Helvetica 9 point type viewed on computer monitors.

In each case, as Tables 2 and 3 indicate, the null can not be rejected. It appears there is no significant legibility difference for Sample A or Sample B based on size of type within the sample.

Table 2 Sample A • Helvetica Normal Type by Size				n=54
Type Style	Mean # words	T-test p-value	Significant	$p=0.05$
Helvetica 12	1.722	p=0.920	no	
Helvetica 10	1.704	p=0.840	no	
Helvetica 9	1.74	p=0.920	no	
Helvetica 12	1.722			

Table 3 Sample B • Helvetica Normal Type by Size				n=53
Type Style	Mean # words	T-test p-value	Significant	p=0.05
Helvetica 12	1.872	p=0.730	no	
Helvetica 10	1.910	p=0.480	no	
Helvetica 9	2.08	p=0.230	no	
Helvetica 12	1.872			

Comparison of Helvetica Bold Type by Size

Null Hypotheses (NH)

- NH4: There is no significant difference between Helvetica Bold 12 point and Helvetica Bold 10 point type viewed on computer monitors.
- NH5: There is no significant difference between Helvetica Bold 10 point and Helvetica Bold 9 point type viewed on computer monitors.
- NH6: There is no significant difference between Helvetica Bold 12 point and Helvetica Bold 9 point type viewed on computer monitors.

Results are mixed. Although Sample A and Sample B did not show significant differences in overall performance for bold, there were differences between samples on legibility of bold type in various sizes. As Tables 4 and 5 indicate, there is no significant legibility difference for Sample A based on size of type for Helvetica Bold.

For the replication in Sample B, however, Helvetica 12 Bold has a significantly higher number of words read than Helvetica 10 Bold and Helvetica 9 Bold and we must reject the null hypothesis that there is no difference between the means for these comparisons.

For Sample A we must accept the null hypotheses and there is no significant difference in legibility based on size for bold face. Sample B gives the same result with one exception. This replication shows we should reject the null and shows that 12 point bold was significantly more legible than 10 point bold or 9 point bold.

Given the mixed results, a closer look at the numbers reveals that the trends are the same for both samples and the mean numbers between the 12 point and 9 point are very consistent. If the two samples are pooled, the T-test results show a significant difference ($n = 107$, Helvetica 12 Bold mean = 2.168; Helvetica 9 Bold mean = 1.800 ; $p=0.005$).

Table 4 Sample A • Helvetica Bold Type			n=54
Type Style	Mean # words	T-test p-value	Significant p=0.05
Helvetica 12 Bold	2.130		
		$p=0.920$	no
Helvetica 10 Bold	2.11		
		$p=0.078$	no
Helvetica 9 Bold	1.790		
		$p=0.052$	no
Helvetica 12 Bold	2.13		

Table 5 Sample B • Helvetica Bold Type			n=53
Type Style	Mean # words	T-test p-value	Significant p=0.05
Helvetica 12 Bold	2.208		
		$p=0.023$	yes
Helvetica 10 Bold	1.830		
		$p=0.92$	no
Helvetica 9 Bold	1.81		
		$p=0.047$	yes
Helvetica 12 Bold	2.208		

Comparison of Helvetica Italic Type by Size

Null Hypotheses (NH)

- NH7: There is no significant difference between Helvetica Italic 12 point and Helvetica Italic 10 point type viewed on computer monitors.
- NH8: There is no significant difference between Helvetica Italic 10 point and Helvetica Italic 9 point type viewed on computer monitors.
- NH9: There is no significant difference between Helvetica Italic 12 point and Helvetica Italic 9 point type viewed on computer monitors.

Results are mixed. Recall Sample A and Sample B did show significant differences in overall performance for Italic with Sample A scoring significantly ($p=0.01$) lower than Sample B for Helvetica 12 Italic and Helvetica 10 Italic.

First note that number of words recognized is very low and in both samples 12 point did better than 10 point which did better than 9 point. As Tables 6 and 7 indicate, there is no significant legibility difference for Sample A or Sample B between Helvetica 12 Italic and Helvetica 10 Italic.

For Helvetica 10 Italic versus Helvetica 9 Italic, Sample A showed no significant difference and Sample B did show a significant difference. Pooling the data shows 10 point Italic significantly better than 9 point although caution must be used as Samples A and B did show a significant difference between samples that may be due to the phrase or words being used ($n=107$, Helvetica 10 Italic mean = 1.110; Helvetica 9 Italic mean = 0.664 ; $p=0.000$).

For Helvetica 12 Italic versus 9 Italic, both samples showed Helvetica 12 Italic with a significantly higher mean number of words recognized.

Based on this analysis, the null is not rejected for NH 7 and is rejected for NH 8 and 9.

Table 6 Sample A Helvetica Italic Type			n=54
Type Style	Mean # words	T-test p-value	Significant p=0.05
Helvetica 12 Italic	0.93		
		p=0.093	no
Helvetica 10 Italic	0.61		
		p=0.350	no
Helvetica 9 Italic	0.463		
		p=0.017	yes
Helvetica 12 Italic	0.93		

Table 7 Sample B Helvetica Italic Type			n=53
Type Style	Mean # words	T-test p-value	Significant p=0.05
Helvetica 12 Italic	1.830		
		p=0.27	no
Helvetica 10 Italic	1.62		
		p=0.000	yes
Helvetica 9 Italic	0.856		
		p=0.000	yes
Helvetica 12 Italic	1.830		

Comparison of Helvetica 12 Point Type by Face

Null Hypotheses (NH)

NH10: There is no significant difference between Helvetica 12 point and Helvetica 12 point bold type viewed on computer monitors.

NH11: There is no significant difference between Helvetica 12 point bold and Helvetica 12 point Italic type viewed on computer monitors.

NH12: There is no significant difference between Helvetica Italic 12 point and Helvetica 12 point type viewed on computer monitors.

Results are fairly consistent between Sample A and Sample B. As Tables 8 and 9 indicate, there are significant legibility differences in 12 point type based on differences in bold and italic face.

For Sample A, Helvetica 12 Bold had a significantly higher number of words recognized than 12 Normal. Helvetica 12 Bold had a significantly higher number of words recognized than 12 Italic and 12 Normal had a significantly higher number of words recognized than 12 Italic.

For Sample B, Helvetica 12 Bold also had a significantly higher number of words recognized than 12 Normal and Helvetica 12 Bold had a significantly higher number of words recognized than 12 Italic. However, 12 Normal did not have a significantly higher number of words recognized than 12 Italic. Again, recall that this was one of the two groups where significant differences existed between the means of Sample A and Sample B for Helvetica 12 Italic so caution should be used in interpretation.

Pooling the data for Helvetica 12 Normal and 12 Italic does indicate a significant difference ($n=107$; 12 point Normal mean = 1.774; 12 point Italic mean = 1.370 $p=0.005$).

The null is rejected in all three cases.

Table 8 Sample A Helvetica 12 point Type			n=54
Type Style	Mean # words	T-test p-value	Significant p=0.05
Helvetica 12	1.722	p=0.023	yes
Helvetica 12 Bold	2.130	p=0.000	yes
Helvetica 12 Italic	0.930	p=0.000	yes
Helvetica 12	01.722		

Table 9 Sample B Helvetica 12 point Type			n=53
Type Style	Mean # words	T-test p-value	Significant p=0.05
Helvetica 12	1.827	p=0.036	yes
Helvetica 12 Bold	2.208	p=0.029	yes
Helvetica 12 Italic	1.830	p=0.990	no
Helvetica 12	1.827		

Comparison of Helvetica 10 Point Type by Face

Null Hypotheses (NH)

- NH13: There is no significant difference between Helvetica 10 point and Helvetica 10 point bold type viewed on computer monitors.
- NH14: There is no significant difference between Helvetica 10 point bold and Helvetica 10 point Italic type viewed on computer monitors.
- NH15: There is no significant difference between Helvetica Italic 10 point and Helvetica 10 point type viewed on computer monitors.

This grouping is one area where results are very inconsistent between Sample A and Sample B. There were not significant differences between Sample A and Sample B for Normal and Bold faces. However, Sample A scored considerably lower on Normal and higher on Bold than Sample B to result in significant differences between Normal and Bold for Sample A. The Italics face showed significant differences between samples and caution needs to be used in interpretation.

As Tables 10 and 11 indicate, there may be significant legibility differences 10 point type based on differences in bold and italic face.

For Sample A, Helvetica 10 Bold had a significantly higher number of words recognized than 10 Normal. Helvetica 10 Bold had a significantly higher number of words recognized than 10 Italic and 10 Normal had a significantly higher number of words recognized than 10 Italic .

For Sample B, there were no significant differences detected between faces.

Based on these inconsistent findings, a pooled sample was tested (table 12) and shows no significant difference between Helvetica 10 and 10 bold but does indicate significant differences in the other two comparisons. Therefore, Null Hypothesis 13 is accepted, there is no significant difference in means while 14 and 15 are rejected.

Table 10 Sample A • Helvetica 10 point Type			n=54
Type Style	Mean # words	T-test p-value	Significant p=0.05
Helvetica 10	1.700		
Helvetica 10 Bold	2.111	p=0.035	yes
Helvetica 10 Italic	0.611	p=0.000	yes
Helvetica 10	1.700	p=0.000	yes

Table 11 Sample B • Helvetica 10 point Type			n=53
Type Style	Mean # words	T-test p-value	Significant p=0.05
Helvetica 10	1.910		
Helvetica 10 Bold	1.830	p=0.720	no
Helvetica 10 Italic	1.620	p=0.260	no
Helvetica 10	1.910	p=0.220	no

Table 12 Pooled Data • Helvetica 10 point Type n=107			
Type Style	Mean # words	T-test p-value	Significant p=0.05
Helvetica 10	1.800		
		p=0.24	no
Helvetica 10 Bold	1.972		
		p=0.000	yes
Helvetica 10 Italic	1.110		
		p=0.000	yes
Helvetica 10	1.800		

Comparison of Helvetica 9 Point Type by Face

Null Hypotheses (NH)

NH16: There is no significant difference between Helvetica 9 point and Helvetica 9 point bold type viewed on computer monitors.

NH17: There is no significant difference between Helvetica 9 point bold and Helvetica 9 point Italic type viewed on computer monitors.

NH18: There is no significant difference between Helvetica Italic 9 point and Helvetica 9 point type viewed on computer monitors.

As Tables 13 and 14 indicate, there are consistent findings between Sample A and Sample B. For both samples tested, Helvetica 9 Normal and Helvetica 9 Bold showed no significant differences in legibility. Helvetica 9 Bold had a significantly higher number of words recognized than 9 Italic and 9 Normal had a significantly higher number of words recognized than 9 Italic .

Therefore, we must accept Null Hypothesis 16 and reject numbers 17 and 18.

Table 13 Sample A • Helvetica 9 point Type n=54			
Type Style	Mean # words	T-test p-value	Significant p=0.05
Helvetica 9	1.741		
		p=0.760	no
Helvetica 9 Bold	1.792		
		p=0.000	yes
Helvetica 9 Italic	0.463		
		p=0.000	yes
Helvetica 9	1.741		

Table 14 Sample B • Helvetica 9 point Type			n=53
Type Style	Mean # words	T-test p-value	Significant p=0.05
Helvetica 9	2.080	p=0.240	no
Helvetica 9 Bold	1.810		
Helvetica 9 Italic	0.868	p=0.000	yes
Helvetica 9	2.080	p=0.000	yes

Rankings by Mean of Words Recognized

In Table 14 each type size and face is ranked by mean number of words recognized from highest to lowest for each of the two samples.

TABLE 15	Sample A	Sample B
Type Style	Mean # words	Mean # words
Helvetica 12 Bold	2.130	Helvetica 12 Bold 2.208
Helvetica 10 Bold	2.110	Helvetica 9 2.080
Helvetica 9 Bold	1.790	Helvetica 10 1.910
Helvetica 9	1.740	Helvetica 12 Italic 1.830
Helvetica 12	1.720	Helvetica 10 Bold 1.830
Helvetica 10	1.700	Helvetica 12 1.827
Helvetica 12 Italic	0.930	Helvetica 9 Bold 1.810
Helvetica 10 Italic	0.610	Helvetica 10 Italic 1.620
Helvetica 9 Italic	0.460	Helvetica 9 Italic 0.868

Note: For each sample a difference of approximately 0.38 to 0.40 results in a significant difference of the means at the 0.05 level.

Since there were no significant differences for 7 or the 9 cells in Sample A and Sample B, pooling the data indicates an overall ranking for the Helvetica type faces and sizes. Caution should be used for the 12 point italic and 10 point italic figures as they did show significant differences between samples.

Table 15		n=107
Type Style	Mean # words	
Helvetica 12 Bold	2.170	
Helvetica 10 Bold	1.970	
Helvetica 9	1.910	
Helvetica 10	1.805	
Helvetica 9 Bold	1.800	
Helvetica 12	1.775	
Helvetica 12 Italic	1.380	
Helvetica 10 Italic	1.120	
Helvetica 9 Italic	0.670	

The first three faces do not show a significant difference from the top performer, Helvetica 12 Bold. Helvetica 10 does have a significantly lower mean than 12 Bold.

The second series does not show a significant difference from the top performer in the section, Helvetica 10.

The third tier does not show a significant difference.

Helvetica 9 Italic stands alone as least legible.

Discussion

Note: For this section, Appendix One will be a useful reference as it shows each type face reproduced at screen resolution and enlarged 200% to show pixel detail.)

Helvetica is a relatively light weight type face using thin, even strokes for each letter. On the computer screen at 72 pixels per inch, this translates to each stroke being one pixel wide. This remains consistent at 12, 10 and 9 point sizes. Therefore, the variability in size comes not in stroke weight, but in the size of the letter forms through x-height and length of ascenders and descenders.

For 12 point type (since there are 72 points per inch and 72 pixels per inch a 12 point face will have 12 pixels height, 10 point = 10 pixels, etc.) the X-height is 7 pixels with ascenders 2 pixels and descenders 3 pixels. The longer descenders give good definition to the "g" and "j".

For 10 point, the X-height is 6, with ascenders and descenders each having 2 pixels. This decrease in descender length hurts the "j" in particular, but other descenders retain their form well. Two pixels is really the minimum for ascenders and descenders as dotted letters need one pixel for the dot and one for the space. Losing an additional pixel on the descenders would fill in the "g" below the line, however this could be considered relatively minor.

However, at 9 point with an X-height of only 5 pixels, letters begin to fill in (notably the "m" and the "w") and a smaller face is not practical on monitors with a 72 pixel per inch resolution.

Helvetica has a very open, rounded letter form that holds the open spaces in letters such as "a" and "e" very well even in smaller sizes. This may help explain why there are no significant differences in legibility of the letters and the number of words recognized between Helvetica 12, 10 and 9 point type faces. Caution should be used here, however, as this research is not measuring readability which would also need to take into consideration variables such as line length and leading.

Helvetica Bold appears to significantly increase legibility for 12 point, may significantly increase legibility (samples differ) for 10 point type and does not decrease legibility at any of the 3 sizes tested. Since Helvetica is a light face, using the bold feature doubles the vertical stroke width to two pixels while leaving the horizontal strokes at one pixel. The effect is to help smooth out thin or jagged edges, especially on letters using slanting strokes such as "v" and "x" and adds weight to rounded forms as well. This provides added visual information for the eye. The open letter forms of Helvetica do not fill in with the additional pixels at the 12 and 10 point sizes and fill very little even at 9 point.

Helvetica Italic should be used with extreme caution. Using bold for emphasis would be much better as it scored significantly higher in both individual samples and in the pooled comparisons.

While results for 10 point are also not as clear, it would be consistent with other findings to use italics with extreme caution with bold being preferred for emphasis.

The results indicate that Helvetica 9 point italic type should not be used as legibility was very low. This face is very light and slanting the letters severely deforms the letter on the monitor.

From this study, we can answer the research questions posed previously:

- 1) There are no significant differences for legibility of Helvetica 12, 10 and 9 point type in the normal face at screen resolutions.
- 2) Helvetica Bold significantly increases the legibility of the type and at the least does not decrease the legibility.
- 3) Helvetica Italics should be used with extreme caution at larger sizes and avoided at smaller sizes.

Appendix One

Figure One • Helvetica alphabets produced and copied at screen resolution of 72 pixels per inch.

abcdefghijklmnopqrstuvwxyz • Helvetica 12
abcdefghijklmnopqrstuvwxyz • Helvetica 12 Bold
abcdefghijklmnopqrstuvwxyz • Helvetica 12 Italic

abcdefghijklmnopqrstuvwxyz • Helvetica 10
abcdefghijklmnopqrstuvwxyz • Helvetica 10 Bold
abcdefghijklmnopqrstuvwxyz • Helvetica 10 Italic

abcdefghijklmnopqrstuvwxyz • Helvetica 9
abcdefghijklmnopqrstuvwxyz • Helvetica 9 Bold
abcdefghijklmnopqrstuvwxyz • Helvetica 9 Italic

Figure Two • Helvetica alphabets produced and copied at screen resolution of 72 pixels per inch and enlarged 200% to show pixel detail.

abcdefghijklmnopqrstuvwxyz • Helvetica 12
abcdefghijklmnopqrstuvwxyz • Helvetica 12 Bold
abcdefghijklmnopqrstuvwxyz • Helvetica 12 Italic

abcdefghijklmnopqrstuvwxyz • Helvetica 10
abcdefghijklmnopqrstuvwxyz • Helvetica 10 Bold
abcdefghijklmnopqrstuvwxyz • Helvetica 10 Italic

abcdefghijklmnopqrstuvwxyz • Helvetica 9
abcdefghijklmnopqrstuvwxyz • Helvetica 9 Bold
abcdefghijklmnopqrstuvwxyz • Helvetica 9 Italic

SOURCES

Felici, J. (1987). Desktop Publishing Skills : A Primer for Typesetting with Computers and Laser Printers. Reading MA: Addison-Wesley.

Felici, J. The Desktop Style Guide. Bantam, ITC.

Felici, J. (1996, March). A Bitmap a Day Keeps the Eyedoctor Away [Internet Document]. Internet locator: <http://www.will-haris.com/typofeli.htm>.

McVey, G.F., & Weigeshaus, C.J. (1973) Studies in Legibility. Rochester, NY: The Center for Visual Literacy, University of Rochester.

Petzold, C. (1992, November 24). Using text and fonts under the OS/2 2.0 Presentation Manager. PC Magazine, Vol. 11, no.20. p. 453ff.

Tinke K. (1996, March/April). Taking It In. Adobe Magazine, pp. 40-45

Tinker, M.A. (1963). Legibility of Print. Ames IA: Iowa State University Press

Will-Harris, D. (1996, March). The Best Faces For the Screen. [Internet Document]. Internet locator: <http://www.will-haris.com/typoscrn.htm>.

Matrix News. (1995). Second TIC/MIDS Internet Demographic Survey. [Internet Document]. Internet locator: <http://www.mids.org.1995>.